DOCUMENT RESUME

ED 216 863 SE 037 270

AUTHOR DeLuca, Frederick P.
TITLE Measurement and Analysis of Logical Thinking. Final Technical Report.
INSTITUTION Iowa State Univ. of Science and Technology, Ames.
SPONS AGENCY National Science Foundation, Washington, D.C.
PUB DATE 26 Mar 80
GRANT SED-77-18654
NOTE 20p.

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *College Science; College Students; *Computer Oriented Programs; Developmental Stages; Higher Education; *Logical Thinking; *Microcomputers; Physical Sciences; Science Education; Science Instruction; Sex Fairness; Social Psychology; Student Characteristics; Teaching Methods; *Testing
IDENTIFIERS *Piagetian Tasks; *Science Education Research

ABSTRACT

Four Piagetian tasks (bending rods, chemical combinations, balanced beams and lights/switches) were programmed on a microcomputer system to rectify perceived deficiencies in the tasks. These deficiencies included misleading perceptual clues, bias against females, familarity with content and task, and high cost of administration and data collection. A microcomputer system for recording and measuring logical thinking ability was developed and tested in experiments designed to study: (1) the ability of subjects (N=394) to demonstrate logical thinking with two kinds of content (physical science and social-psychology); (2) patterns of logical thinking under two kinds of instruction (global and differentiated); and (3) automation of data collection. Findings from these studies indicate that: (1) the microcomputer is a useful tool for studying logical thinking; (2) teacher-learning variables can produce significant gains at the concrete operational level for college students, but gains at the formal operational level are more difficult; and (3) successful problem solvers tend to use relatively few, but highly efficient patterns. Nine advantages (including a significant reduction of task bias against female subjects) and four disadvantages of the microcomputer system are listed and discussed. (Author/JN)

DATA ON SCIENTIFIC COLLABORATORS

Final Project Report for Grant SED77-18654

March 26, 1980

1. Principal Investigator

   Dr. Frederick P. DeLuca

   Associate Professor of Earth Science & Science Education

   Iowa State University

   Ames, Iowa 50011

2. Faculty Associate

   Dr. Donald C. Scouten

   Associate Professor of Electrical Engineering

3. Unpaid Consultant

   Dr. Warren D. Dolphin

   Professor of Zoology

4. Research Assistants

   Mrs. Anita Cody

   Graduate Student

   Mrs. Evelyn Beavers

   Graduate Student

   Mrs. Eden Schmitt

   Graduate Student

FINAL TECHNICAL REPORT

for NSF Grant SED77-18654

MEASUREMENT AND ANALYSIS OF LOGICAL THINKING

Principal Investigator:

Dr. Frederick P. DeLuca

Department of Earth Sciences

253 Science I

Iowa State University

Ames, Iowa 50011

Submitted to NSF on March 26, 1980

Summary of What was Proposed

The principal investigator indicated that Piagetian tasks suffered from misleading perceptual cues, bias against females, familiarity with content and task, and the high cost of administration and data collection. These deficiencies limited the effectiveness of Piagetian tasks as vehicles for measuring and studying logical thinking.

It was proposed that the deficiencies could be rectified and Piagetian tasks improved upon by programming the tasks on a microcomputer system. The proposal called for the development of a microcomputer system and programmed tasks which were to be tested in three experiments designed to study: 1) the ability of subjects to demonstrate logical thinking with different kinds of content (physical science and social-psychology), 2) patterns of logical thinking under two kinds of instruction (global and differentiated), and, 3) automation of data collection.

The primary goal of the project was to develop instrumentation that would permit more accurate observation and measurement of logical thinking than was provided by Piagetian tasks. This goal was differentiated into several questions:

1. Can Piagetian tasks be programmed so that their deficiencies are eliminated?
2. Can the interaction of the subject with the programmed tasks be monitored and recorded for subsequent analysis?
3. Can the recorded patterns of interactions between subjects and programmed tasks be indicative of the quality and level of logical thinking?
4. Can the data collection process be automated?
5. Do subjects score equally well when the task and level of logical thinking are held constant and content (physical science and social-psychology) are varied?
6. Do subjects score equally well when the task and level of logical thinking are held constant and task instructions (global and differentiated) are varied?

Summary of What was Done

All work was completed essentially as proposed. The only changes occurred in experiment #3, where the primary objective, automation of data collection, remained the same but the sample was changed to gain knowledge of the influence of gender and major on the problem solving abilities of seniors in physical sciences (engineering, physics, chemistry), mathematics, and English-speech. This decision was based on the results of experiments #1 and #2, and the desire to obtain a sample that was more polarized in terms of problem solving abilities.

Development of Hardware and Software

Hardware

The original goal was to custom-build a microcomputer system with a combination of off-the-shelf components and some individually designed electronic circuit boards. However, in the intervening time between proposal and funding, some microcomputers and floppy disk drives were introduced at prices that made it more economical to purchase fully assembled commercial units. The initial system consisted of an Apple II microcomputer with 32k bytes of memory, a Ball Brothers 9-inch black and white video monitor, and a PERSCI 8-inch dual floppy disk drive and intelligent controller. Although the Apple II contained all of the necessary interface hardware (alphanumeric TV and high resolution TV drivers, cassette interface, and a keyboard), it lacked a real time clock. This deficiency was corrected by mounting a 6522 Versatile Interface Adapter (VIA) on a circuit board and plugging it into one of the eight peripheral female connectors which provided access to the microcomputer's bus. The VIA contained both a counter and the necessary circuitry to access the bus when addressed. The counter, which was incremented 10 times per second to provide a 0.1 second resolution, utilized the normal memory space and was accessed directly from the Apple II Basic program by use of PEEK and POKE commands.

Despite the Apple's capacity to display 23 lines of program on command, the long and complex Basic programs were difficult to debug and modify without a hard copy. The addition of a Centronics 779 line printed resulted in a much higher programming efficiency and was well worth the extra cost. Use of the VIA proved fortuitous in that it also provided an output port to drive the line printer and eliminated the need for an additional interfacing circuit board.

Software

Five programs totaling about 80k bytes of memory were written. Except for some minor subroutines for the high resolution graphics and data recording which required machine language, all the programs were written in Apple II Basic. This proved to be a convenient and self-documenting format that facilitated program modifications and corrections.

Three of the programs were based on the Piagetian tasks known as the Separation of Variables (or Bending Rods), Chemical Combinations, and the Balanced Beam (1). The fourth program was based on a Piagetian-type task entitled Switches and Lights as described by Davis (2). The fifth program was based on the Bem Sex Role Inventory (BSRI), a test for characterizing a subject as masculine, feminine, or androgynous (3). Some of the experiments required that the original programs for the first four tasks be modified to simulate two different types of content and two different types of instruction. Thus there were four different versions of each of the four tasks: two versions involving different subject matter content (physical science and social psychology) and two versions involving different instructions (differentiated and global). The same program for the BSRI test was used in all three experiments.

All programs were designed to promote maximum interaction between the system and the subject because the intent was to collect data on the process (form and quality of the interaction) as well as the product of logical thinking. In each program the subject was introduced to a specific task, given directions, and provided with opportunities to interact with the program. In most cases, depending on the experiment and the type of task, subjects were asked to experiment, to learn as much as possible about the variables, to solve the problem, or to show proof of his or her understanding. At other times, subjects were asked rather direct questions that required them to make a selection from two or more responses.

Operating Procedures

The general operating procedure for the collection of data was as follows. Two graduate students scheduled subjects for interaction with the system and served as monitors during the collection of data. A graduate student started the typical day of data collection by inserting the program and data disks into their respective slots of the floppy disk drive and loading the first program into the Apple II. At this point the video screen read, "I AM READY. PRESS THE SPACEBAR TO BEGIN." The subject pressed the spacebar and proceeded through an introductory program which was designed to relax the subject and obtain his or her coded identification number, first name, and last initial. The information was temporarily stored in the Apple II. Upon the response to the last question of the introductory program, the information was automatically transferred to the data disk, the introductory program was "erased" from the Apple II, the next program was automatically loaded into the Apple II, and directions appeared on the screen.

As the subject interacted with the program, all depressions of keys and the time between depressions were recorded and stored in the Apple II. The subject's response to the last phase of the task initiated transfer of the data disk and automatically loaded the next program. This process continued until the subject completed the last program and the screen read, "THANK YOU. THIS CONCLUDES THE PROGRAM." After a delay of sufficient length to allow the subject to leave, the screen returned to the beginning of the introductory program and the screen read, "I AM READY. PRESS THE SPACEBAR TO BEGIN."

After each data disk became full, it was taken to the Iowa State University Computer Center and interfaced with the IBM/370 in order to transfer the data onto tape and obtain hard copies of the data. This made it possible to perform any type of data manipulation and analyses that were available at the computer center.

Sample

The total sample consisted of 394 undergraduate students. Experiment #1 utilized a stratified (male, female, science major, non-science major) random sample of 120 students (60 males and 60 females) from introductory biology courses. Experiment #2 utilized a similar sample of 142 students (70 males, 72 females). Experiment #3 utilized a stratified (male, female, physical sciences, mathematics, English-speech) random sample of 132 seniors (68 males, 64 females).

Type of Variables and Data Analysis

The intent was to measure all possible variables that could provide information on the process and the product of the student's interaction with each task. Although some variables were unique to specific tasks, the following list is representative of the variables measured: practice tries, practice time, problem tries, problem time, errors, recall of key combinations, chunks of information, proofs, and level of solutions.

Data were analyzed by way of analysis of variance, chi square, correlation matrices, cluster analysis and special sorting programs for the identification of patterns. In addition, a great deal of time and effort was spent in examining individual sequences of responses in an attempt to gain insights as to why certain students used certain patterns. (This effort is continuing.)

Experiment #1: Physical Science and Social-Psychology

(A manuscript on this experiment is in preparation)

In this experiment there were two versions of each of the four Piagetian tasks. In one version the task variables were given physical science labels, whereas, in the alternate version the variables were given social-psychology labels. This approach made it possible to present the same task in the context of physical science and social-psychology content, respectively. One-half of the sample worked with the physical science content, and the other half of the sample worked with the social-psychology content.

Results of Experiment #1

Analysis of data on the balanced beam failed to yield any significant differences. The type of content had no significant (.05) effect on any of the variables.

For the switches and lights task, significant (.05) differences on practice tries, practice time, and solution were related to major, gender x major, and major, respectively. The type of content had no significant (.05) effect on any of the variables.

For the chemical combination task, significant (.05) F values indicated that students who worked with physical science content committed fewer errors (repeats and omissions) and had greater recall of key combinations than did students who worked the tasks in the context of social-psychology content. A significant (.05) interaction (content x major) indicated that science majors used fewer tries under physical science content and more tries under social-psychology content as compared to non-science majors. There were no significant differences on measures of level of solution.

For the bending rods task, significant (.05) F values indicated that students who worked with physical science content took significantly less time and scored significantly higher on measures of level of solution. It should be noted, however, that the transformation of this task into an equivalent social-psychology content task was complex, and as a result the social-psychology version was much more difficult.

Interpretation of Experiment #1

If the bending rods task is disregarded on the basis that the two versions were not equivalent, three tasks remain. The balanced beam, and switches and lights tasks were both unaffected by content differences. In the chemical combinations task, the number of tries (content x major), number of errors and recall of key combinations were all significantly (.05) affected by task content. It is interesting and important to note that all of these variables are characteristic of the concrete phase of the chemical combination task. The level of solution (isolation of the effect of each of the four "chemicals") constitutes the formal operational phase of the task, and it was not significantly (.05) influenced by task content. Thus, the interpretation is that in some cases different content will not affect student performance on a task, and in some cases the concrete operational phase of the task will be significantly affected, but it is less likely that the formal operational phase of a task will be affected significantly.

Expe_iment #2: Global and Differentiated Instructions

(A manuscript on this experiment is forthcoming.)

The four Piagetian tasks which were used in experiment #1 were also used in experiment #2 but with modifications. The chemical combinations task was modified only slightly to improve interaction with the students. The switches and lights task remained essentially the same, but three switches and three subtasks were added to make the task more challenging. The balanced beam was reprogrammed so its operation was based on the formula that torque is equal to weight divided by distance from the fulcrum. This was done to prevent task familiarity from influencing student performance. The bending rods task underwent the most modification in that the format was changed from a diagram of six rods to a chart of five systems (rods) each containing three fixed variables and two changeable variables. The amount of bending was referred to as level of reaction and was indicated in digital form for each change in variable.

In this experiment there were two versions of each of the four Piagetian tasks. In one version the task instructions were represented in global form. For example, "Experiment and learn as much as you can about the system." In the alternate version task instructions were differentiated into a list of specific steps. For example, 1) Find which switches control certain lights, 2) Find how the switches and lights interact, 3) Look for patterns between switches and lights. One-half of the sample worked with the global instructions, and the other half of the sample worked with the differentiated instructions.

Results and Interpretation of Experiment #2

Analysis of the data yielded many significant (.05) main effects and interactions on process variables, however, there were no significant (.05) differences on measures of the solutions. It appears that differences in instructions affected the concrete operational phases of the tasks but did not affect significantly (.05) the formal operational phases (solutions). The interpretation here, as in experiment #1, is that it is much easier to influence significantly (.05) a student's performance on the concrete operational phase of a task than it is on the formal operational phase of the task.

In regard to the analysis of patterns, it appears that patterns are influenced by many variables, major, gender, task, task content, and task instructions. Many different patterns can lead to a successful solution; however, similar patterns do not necessarily lead to similar solutions. When the range of patterns are considered, it appears that the best problem solvers tend to use patterns which require low memory loads.

Experiment #3: Automation of Data Collection

(A manuscript on this experiment is forthcoming)

The prime objective of this experiment was to determine whether the data collection process could be fully automated. A second objective was to obtain data that would be representative of successful problem solvers and unsuccessful problem solvers, so that the analysis of patterns could be facilitated. A third objective was to investigate the relationship between scores on the Bem Sex Role Inventory and college major. It was reasoned that a sample of mostly seniors (and some juniors) in physical science (engineering, physics and chemistry), mathematics, and English-speech, would be best for meeting the above objectives.

The four Piagetian tasks used in this experiment were essentially the same as the differentiated tasks which were used in experiment #2. All of the tasks, however, were modified to provide greater clarity and ease of program-flow. All students worked with the same sequence of tasks. When each student arrived for testing he or she was assigned an identification number which he or she typed on the keyboard. There was no further communication between the student and the monitor who had assigned the identification number.

Results and Interpretation of Experiment #3

Examination of the recorded data and comparison with data from earlier experiments indicated that 98 percent of the data was classified as good data. The other two percent was considered unuseable because data were missing. Reports from the monitors indicated that the students did not experience any difficulty in understanding and following the programs. These results support the view that data can be collected on an entire battery of tasks without the interaction of an expensive examiner. This means that the system can be placed in a library or a media resources room, and data can be collected at a very low cost.

Analysis of variance indicated that a total of 12 significant (.05) F values on three tasks were related to college major. Only the balanced beam failed to produce any significant differences due to college major. Thus, these data are well suited for analysis of patterns. (Work is still progressing in this area.)

Scores on the Bem Sex Role Inventory (BSRI) yielded significant (.05) F values on gender, major, and gender x major. The data indicates that there is a significant difference between groups of college majors in terms of how they view themselves on the feminine-masculine spectrum of roles. It appears, however, that these differences are unrelated to problem solving abilities. (Work is still progressing in this area.)

Advantages of the Microcomputer System

Compared to the traditional form of Piagetian tasks, where an examiner tests a single subject at a time, the microcomputer system described above proved to have many advantages.

1. Standardization of Testing Procedures. The microcomputer system insures that the testing procedures will be exactly the same for all subjects and thereby eliminates the concern for biases that may occur between examiners and within a single examiner when testing subjects in the traditional manner. This also means that there will be a much better basis for making comparisons between research results from several research centers.

2. Control of Misleading Perceptual Cues. The microcomputer system permits precise control of every bit of information presented to each subject. Thus misleading perceptual cues can be eliminated or they can be introduced under controlled conditions. In this way the relationship between misleading perceptual cues and success in problem solving can be studied and better understood.

3. Control of Task Familiarity. Many high school and college students have studied the balanced beam, so one should expect that such familiarity would influence their performance on the task. In two of the experiments in this study, the balanced beam was programmed to behave very differently from an ordinary balanced beam. Mysterious matter called "stuff" exerted an upward force on the beam and "torque" was inversely proportional to the distance away from the fulcrum rather than directly proportional as is the case with the regular balanced beam. Because no one had ever encountered such a balanced beam in the real world, subjects could not have been familiar with it. Ability to solve the problem depended on the subject's ability to experiment, analyze, and integrate information rather than on mere recall of a familiar experience.

4. Control of Content Bias. In an alternate version of each of the four programmed tasks, the physical science variables were assigned social-psychology labels. For example, the chemicals in the chemical combinations task were assigned names of people, and the chemical reaction was referred to as a reaction between people. No reference was made to chemicals. The same level of logical thinking was required to solve the problem, but the content of the task was different. The claim that Piagetian tasks are biased in favor of subjects who are familiar with physical science concepts did not apply in this case. This approach made it possible to keep the required level of logical thinking constant and to study the effects of different content on the performances of individuals.

5. Increased Comprehensiveness and Accuracy in Data Collection. The microcomputer system records the value of each variable as the key for that variable is depressed, so the recorded data provide the values of all variables

for any given time during the subject's interaction with the system. Such a collection of data would be very difficult to obtain by way of the traditional Piagetian tasks. Moreover, the electronic method of recording data and transferring it to IBM 370 compatible tape reduces the chance of human error that exists when data are recorded by hand and then keypunched on computer cards.

6. Reduction of Task Bias Against Female Subjects. Of the approximately 100 tests for significant (.05) differences between males and females, 4 significant tests were obtained on the measures of practice time and time required to complete the tasks. No significant differences, based on gender alone, were found on measures of success in solving the tasks. These were, however, 10 significant interactions between gender and major, gender and task content, and gender major and task content on various subtask elements. Compared to the results obtained by other researchers (4, 5), it appears that the programmed tasks reduced bias against female subjects. This view, however, must be tempered with caution because the sample in this study included only college students.

7. Lower Cost of Data Collection and Analysis. The cost of data collection is reduced because subjects can proceed through the entire set of programmed tasks without the guidance of a highly paid, skilled examiner. If a monitor is desired to insure proper identification of subjects and protection of the hardware, he or she could be a relatively low paid, unskilled person. Also, money can be saved because the data need not be tabulated by hand nor keypunched as is commonly the case when data are collected by hand.

8. Accurate Measurement of Real Time. Every time a key is depressed, the elapsed time since the last key depression is automatically recorded in tenths of a second. Thus the rates of progress through any segment of the task, as well as the total time required to complete the task, is known for all subjects.

9. The System is Transportable. In the not too distant past the above advantages could only be achieved by transporting students to a fixed based computer facility. The microcomputer system can be transported wherever the subjects are located, be it in a dormitory, library, student union, or public school.

Disadvantages of the Microcomputer System

Compared to the traditional form of Piagetian tasks, the microcomputer system does have some disadvantages.

1. High Initial Cost of Hardware. The materials needed to administer a battery of Piagetian tasks in the traditional manner cost less than fifty dollars. The cost of hardware for the microcomputer system is between five and six thousand dollars. Although the cost of hardware continues to drop, its cost will always be more expensive than the simple laboratory materials used in Piagetian tasks. (In the long run, however, the automation of data collection through use of the microcomputer system will result in lower costs per subject.)

2. Writing Computer Programs is Very Time Consuming. A great deal of time is required to write a good program. It is not so much a matter of debugging a program; much more time is required to write and rewrite the

program until it represents a good task in terms of logical thinking, and it is easily followed by the type of subjects who are to be tested. In the future this problem could be diminished if researchers share their programs with each other.

3. Repair of Hardware Requires Expert Help. The materials used in the traditional Piagetian tasks require only minor attention to repairs, so the researcher can be in full control of the entire research project. The use of electronic and electro-mechanical hardware means that equipment will fail and repairs will be necessary. When this occurs, the researchers must rely on outside expert help that may result in long delays and added expense.

The second author of this article is an electronic engineer. His expertise proved to be critical to the success of this project. Several failures in equipment were quickly repaired and time lost was kept to a minimum. The PERSCI disk, however, developed some problems that could not be repaired locally because specialized alignment equipment was needed. The disk drive was sent to California for repairs on two occasions that resulted in delays of two and six weeks, respectively.

4. Programmed Tasks are Less Flexible. The programmed tasks are presented in exactly the same way to all subjects. Although this is advantageous in that it forms a common standard, it lacks the capacity to adjust to different subjects and extract insights that are possible with a skilled examiner in the original Piagetian tasks. Perhaps some researchers will opt for a blend between the two forms of tasks and choose to use a skilled examiner as an integral part of the programmed tasks.

Answers to the Original Questions

The primary goal of this project was to develop instrumentation that would permit more accurate observation and measurement of logical thinking than was provided by Piagetian tasks. The original questions and their answers are as follows:

1. Can Piagetian tasks be programmed so their deficiencies are eliminated?

   Answer: Yes, not only can deficiencies be eliminated, but advantages beyond Piagetian tasks can be obtained.

2. Can the interaction of the subject with the programmed tasks be monitored and recorded for subsequent analysis?

   Answer: Yes, definitely.

3. Can the recorded patterns of interactions between subjects and programmed tasks be indicative of the quality and level of logical thinking?

   Answer: In terms of an individual's pattern, the answer is, no. In terms of a group of individuals, the answer is a weak, yes. (Work is continuing in this area.)

4. Can the data collection process be automated?

   Answer: Yes, and at reduced costs over the longrun.

5. Do subjects score equally well when the task and level of logical thinking are held constant and content (physical science and social-psycgology) are varied?

   Answer: If "equally well" refers to the level of solution (formal operational level), the answer is, yes. If "equally well" refers to the concrete operational phase of the task, the answer is both, yes and no, depending on the task.

6. Do subjects score equally well when the task and level of logical thinking are held constant and task instructions (global and differentiated) are varied?

Answer: The answer is similar to that of question #5. Yes, if we are talking about the formal operational level (solutions). No, if we are talking about the concrete operational phase of a task.

References

1. B. Inhelder and J. Piaget, The Growth of Logical Thinking From Childhood to Adolescence, Basic Books, New York, 1958.
2. G. A. Davis, Psychology of Problem Solving, Basic Books, New York, p. 44, 1973.
3. S. L. Bem The Measurement of Psychological Androgyny, Journal of Consulting and Clinical Psychology, 2, pp. 155-162, 1974.
4. A. E. Lawson, Sex Differences in Concrete and Formal Reasoning Ability as Measured by Manipulative Tasks and Written Tasks, Science Education, 3, pp. 397-405, 1975.
5. L. G. Dale, The Growth of Systematic Thinking: Replication and Analysis of Piaget's First Chemical Experiment, Australian Journal of Psychology, 22, pp. 277-286, 1970.